Faculty of Computer Science and Information Technology

Malaysia Ethnicity-based Facial Expression Classification and Emotion Mapping

Izzah Nilamsyukriyah Binti Buang

Master of Science
2021

Malaysia Ethnicity-based Facial Expression Classification and Emotion Mapping

Izzah Nilamsyukriyah Binti Buang

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Computer Science)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2021

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

………………………………

Signature

Name: Izzah Nilamsyukriyah Binti Buang

Matric No.: 17020132

Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

Date:

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and the Most Merciful.

Alhamdulillah, all praises to Allah for the strengths and His blessing in completing this thesis. This thesis has been kept on track and been seen through to completion with the support and encouragement of numerous people including family and friends.

Special appreciation goes to my supervisor, Dr. Hamimah Binti Ujir, for her supervision and constant support, the invaluable help of constructive comments and suggestions throughout the making to the success of this research. My deepest gratitude goes to my beloved parents Mr. Buang Bin Asin and Mrs. Maimunah Binti Buang and also to my brothers for their endless prayers and encouragement. Sincere thanks to my strong supporters and friends for their moral support, kindness and being there along the struggles, also bring joys and laughter. Thanks for the friendship and memories. Finally, thank all those who have helped me directly or indirectly in the successful of the thesis. JazakAllahu khair. May Allah reward you for the good.

ABSTRACT

Human are commonly good in notifying several emotions via facial expression. In daily human communication, it is crucial for each person to be able to convey his emotions and perceive others respectively using speech, facial expressions and body movements. The computer vision experts are continuously learning on how to achieve high performance in analysing faces, especially which occur spontaneously. Malaysian facial database and analysis are still inconspicuous, especially for local ethnicity studies. Hence, this thesis developed MUA Database, the first Malaysian ethnicity facial database, which consists of data from non-actor subjects from 4 local ethnicities that are Chinese, Iban, Indian and Malay. During the data collection, the subjects are encouraged to express facial expressions spontaneously. Facial expressions analyses are done using the database and facial deformation for each ethnicity is evaluated. From the experiments, the performance of HOG, LBP and SIFT are compared for feature extraction, and SVM, Decision Tree and KNN performance are evaluated as classifier. Results show that the combination of HOG features and KNN classifiers are the best pair for ethnic recognition with 96.90% accuracy, whereas HOG features and SVM classifier combination shows the best pair for emotion recognition with 59.10% accuracy. Indian appeared to be the most recognisable among other ethnicities. As for emotion, “happy” appear to be the most conspicuous emotion, whereas “fear” is the least visible among all tested emotion.

Keywords: Ethnic facial analysis, face recognition, face database, ethnicity classification, ethnic face database

Klasifikasi Ekspresi Muka Berdasarkan Etnik Malaysia dan Pemetaan Emosi

ABSTRAK

Manusia lazimnya boleh mengesan beberapa emosi menerusi ekspresi muka dengan mudah. Penting bagi setiap orang untuk menyampaikan dan memahami emosi orang lain melalui nada dan intonasi percakapan, ekspresi muka, dan bahasa badan dalam komunikasi harian. Pakar visi komputer sehingga kini berterusan mengkaji untuk mendapat pencapaian yang tinggi dalam proses menganalisa muka, terutamanya yang timbul secara spontan. Pangkalan data dari Malaysia masih belum kukuh, khususnya dalam kajian mengenai etnik tempatan. Justeru itu, tesis ini memperkenalkan Pangkalan Data MUA, pangkalan data wajah ethnic Malaysia yang pertama, merangkumi data bagi peserta bukan artis daripada 4 etnik utama yang terdiri daripada Cina, Iban, India dan Melayu. Peserta digalakkan untuk mengekspresikan riak muka secara spontan semasa proses pengumpulan data. Analisa bagi ekspresi muka juga dilakukan menggunakan pangkalan data tersebut sekaligus mengkaji dan menilai perubahan wajah pada setiap etnik. Tambahan lagi, melalui ini turut membandingkan prestasi HOG, LBP dan SIFT bagi proses pengekstrakan ciri serta menilai prestasi SVM, Decision Tree dan KNN sebagai pengelas. Hasil menunjukkan ciri HOG dan pengelas KNN merupakan kombinasi terbaik bagi mengenalpasti etnik dengan 96.90% ketepatan, manakala ciri HOG dan pengelas SVM adalah gabungan terbaik bagi mengenalpasti emosi dengan 59.10% ketepatan. Keputusan menunjukkan India merupakan etnik yang paling mudah dikenali berbanding etnik lain. Emosi "Gembira" merupakan emosi yang paling menyerlah, manakala emosi "Takut" paling sukar dikesan.

Kata kunci: *Analisis wajah mengikut etnik, mengenalpasti wajah, pangkalan data muka, pengelasan etnik, pangkalan data muka mengikut etnik.*

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

LIST OF ABBREVIATIONS

AAM	Active Appearance Models
AFEW	Acted Facial Expression in the Wild Database
AI	Artificial Intelligence
ANN	Artificial Neural Network
AOML	Adaptive Online Metric Learning
AU	Action unit
BU-3DFE	Binghamton University 3d Facial Expression
CAS(ME)2	Chinese Academy of Sciences Macro-Expressions and Micro-Expressions Database
CLM	Constrained Local Model
CNN	Convolutional Neural Network
CV	Confidence value
DBN	Dynamic Bayesian Network
DISFA	Denver Intensity of Spontaneous Facial Action
DL	Dictionary Learning
EUC	Euclidean distance metric
FACS	Facial Action Coding System
FEED	Facial Expression and Emotion Database
FLIR	Forward Looking Infrared thermal camera
HCI	Human Computer Interaction
HOF	Histograms of Optic Flow
HOG	Histogram of Oriented Gradient

ICA	Independent Component Analysis
KNN	K-Nearest Neighbor
LBP	Local Binary Pattern
LBP-TOP	Local Binary Pattern histograms from Three Orthogonal Planes
LDA	Linear Discriminant Analysis
LDCRF	Latent-Dynamic Conditional Random Field
LGBP	Local Gabor Binary Pattern
LLE	Locally Linear Embedding
LoG	Laplacian of Gaussian
LPQ	Local Phase Quantization
MFCC	Mel-Frequency Cepstral Coefficient
MFP	Moving Faces and People database
NNC	Nearest Neighbor Classification
nrML	Non-robust version of metric learning
OEE	Other-ethnicity-effect
p.s.d.	Positive semidefinite
PCA	Principal Component Analysis
PLP	Perceptual Linear Prediction
PLS	Partial Least Square regression
POEM	Patterns of Oriented Edge Magnitudes
RASTA	Relative Spectral
RobustML	Robust version of metric learning
SIFT	Scale-invariant Feature Transform
STASM	Stacked Active Shape Model facial landmark

SVM	Support Vector Machine
UNBC-McMaster	Mc Master University and University of Northern British Columbia
USTC-NVIE	University of Science and Technology China Natural Visible and Infrared facial Expression database

CHAPTER 1

INTRODUCTION

1.1 Introduction

Human facial analyses are often in deliberation for various applications nowadays, starting with scientific to industrial communities competing to create consecutive value to human daily life. The utilization of facial analysis is no longer limited to surveillance and security (Chang, 2004; Beghdadi et al., 2018), but also contribute in cognitive studies (Palestra, 2018; Yu, 2018), biometric (McLain and Kefallonitis, 2019; Wouters et al., 2019), virtual reality (Briggs et al., 2018; Souto et al., 2019), multimedia (Tao et al., 2018; Chen & Nahrstedt, 2019), computer entertainment (Braun et al., 2018), and have significant part in artificial intelligence (AI) applications (Sharma et al., 2019; Wang et al., 2019). In order to create machines that can interact efficiently with human, the development of human computer interaction (HCI) nowadays is no longer restricted to constructing the groundwork of user interface. Various applications have included human cognitive and behavioural model that are producing user-centric design concept which linked to human life in order for the interaction to be more effective and efficient (Wang et al., 2019), as for example the progressive development of mobile phone interface (Lalji & Good, 2008). Therefore, the studies of human faces are crucial in order to develop machine that is able to respond to facial information rather than to depend on input commands from the user.

This thesis will focus on the challenges encountered in analysing facial expressions which are still unsolved, measure the accuracy of existing solutions and also suggest solution for the problems. We also focused to bring the problems locally, so that the specific information and studies can be obtained for the selected geographical area.

1.2 Motivation

The field of computer vision have many applications in various ways nowadays, where machine learning is performed to interpret visual information with an intention to improve the quality of life. It helps in detecting and describing knowledge in images, interpreting and even learning solutions by using the computer (Szeliski, 2010). Face analysis is one of essential process of computer vision. The information from this process enable us to learn a lot more about people, from their identity, ethnicity, age, gender, expression and emotions. These practices are then implemented to various applications, such as in HCI, human evaluation and cognitive, security and a lot more.

The most interesting part to analyse from faces is the emotion. Emotion recognition a primary framework for of emotion regulation, which both important in human relationships as the ground in reasoning, problem solving and enhancing cognitive activities (Mayer et al., 2001). This aspect is essential especially for multicultural country like Malaysia as it can be used to improve the intercultural adjustment between ethnicities (Yoo et al., 2006). In order to classify emotion from faces, a lot of complex algorithms will be involved, which some of them will be discussed in Chapter 2. Principally human display their emotion through facial expressions (Hupont et al., 2010). One of major face recognition and classification problem is usually occurred when large target set involved. Thus, specification of ethnicity, age and gender may increase the accuracy of the identification process. Human cognitively perceive ethnicity before age, gender and expression from face (Wang et al., 2019). Therefore, this research will concentrate on ethnicity-based facial expression analysis. Ethnicity defined as group of people in a region with shared culture, language and some with particular skin colour (Braun et al., 2013). Malaysian facial database and analysis are still inconspicuous, especially for local ethnicity studies. Moreover, there is no existing Malaysia ethnics' database and facial emotion analysis available. So, this research will analyse facial expressions of the main ethnicities in Malaysia which are Chinese, India and Malay, and also Iban, which is one of major ethnic in Sarawak, Malaysia.

The computer vision experts are continuously learning on how to achieve higher performance in analysing faces, especially which occur spontaneously (Saha et al., 2019). Since the study of spontaneous face is imperative and complex (Liu & Yin, 2017), as it has more variations in term of physical peculiarity and class compared to the basic facial expression (Ekman, 2001). Generally, psychologist represent facial emotion in discrete division of six universal emotions, which are "happy", "sad", "fear", "anger", "disgust" and "surprise" (Ekman et al., 1999). However, human daily interaction involves broad range and intensity of emotion other than the six universal emotions (Hupont et al., 2010). Therefore, for this research, other non-basic affective states in describing emotion will be identified using emotional mapping.

1.3 Problem Statement

Humans are commonly good in recognizing several emotion facial expressions. "Happiness" and "surprises" can be easily perceived compared to "anger" and "sadness", and even worse for "fear" and "disgust" (Martinez & Du, 2012). This is probably because the emotions of "happiness" and "surprise" are involving wider face transformations than the rest (Saha et al., 2019). In daily human communication, it is crucial for each person to be able to convey his emotions and perceive others respectively using speech, facial expressions and body movements. These processes are usually dynamic and spontaneous.

Spontaneous expressions give a huge challenge as they are exempt from any intentional attempt, not always noticeable and not fully expressed (Ekman, 1997) due to the coexistence of other basic expressions of emotion (Reisenzein et al., 2013). It is more complicated and the changes are more gradually than the acted one, directing to subtle sequence of expressions (Ekman & Friesen, 1976). This causes fuzzy distinction between different emotions. Each individual also portrays their emotions in various manners, which resulting diverse and confusing information of similar emotions due. For example, Tarnowski et al. (2017) have able to recognised discrete basic emotion using facial expression and shows the most recognition mistakes occurs between each pairs of expression, but failed to represent emotion based on the presence of each expression occurred from a spontaneous expression. As mentioned by Ekman and Friesen (1976), the existence of only one specific basic facial expression of emotion is rarely to appear. Thus, the ability to only detect 6 basic emotion facial expressions distinctively are not enough in order to describe spontaneous expressions due to intra-class diversity (Gazizullina & Mazzara, 2019).

An ideal emotion detection system should be able to recognized expressions regardless of their gender, age and ethnicity besides the ability to be consistent through various diversions including illumination, lightening conditions, glasses, facial hair and hairstyles (Sebe et al., 2004). Izard (2009) questioned whether some emotional expressions are universal, which means the individuals of distinct cultures resulting in similar facial muscle movements when expressing some emotion. In developing facial recognition system, ethnic identification helps on identity-related features, and narrows down the search in large database which will increase the search speed and efficiency of the system. This also applies to demographic statistics in many social applications (Lu & Jain, 2004). Due to ambiguous physical and psychological nature of ethnical group, it is a comprehensive problem in differencing emotions using facial expression because of the existence of similar facial features and characteristics among the ethnicities (Wang et al., 2019). Ma et al. (2019) had developed a facial expression database of Chinese ethnicity, but has not mention any distinction of ethnicity for each facial expression. Thus, there is a compelling need to have a facial analysis on different ethnicities (Wang et al., 2019).

1.4 Objective

The primary objectives of this research are as follows:

a. To determine the association between facial deformities and ethnicity in expressing different type of facial expression
b. To classify the affective values of spontaneous expression from MUA Database through emotion mapping

In this research, a Malaysian ethnicity face database named MUA Database is developed, which also include spontaneous facial expression dataset. From this, it should able to classify spontaneous facial expression images of the database through emotion mapping. The hypothesis for this thesis is the ethnicity categorization may improve the facial expression classification using FACS. Other hypothesis is that the deformation of facial expression using FACS can improved ethnicity classification.

1.5 Scope of Research

The main focus of the project is the study of facial expression classification between Malaysian ethnicity and emotion mapping of spontaneous facial expression using Whissell Space. The study is limited to 4 major ethnicities in Malaysia, which are Chinese, Iban, India and Malay. The development of MUA Database include 200 subjects and this research involve data training, testing and analysis from the database. The emotion mapping from the research also only using spontaneous facial expression from the database. In this research will only compare the performance of 3 facial features which are HOG, SIFT and LBP, and 3 classifiers that are SVM, Decision Tree and KNN.

1.6 Significance of the Study

This research helps to investigate the emotion facial recognition and analysis between ethnicity in Malaysia. The findings of this study will directly benefit to improve local face recognition system and intercultural studies, either cognitively or psychologically, between ethnicities in Malaysia. Main contribution of the thesis is the development of the first Malaysian ethnicity database, MUA Database, and facial expression analysis. This research also contributes in publishing a paper that discuss on Malaysian ethnicity recognition (Buang & Ujir, 2019).

1.7 Thesis Outline

The thesis is organized into five chapters. Chapter 1 introduces the research work. Chapter 2 presents the literature review and current research that bring out the problem and restate the need of the study of facial ethnicity, expression and recognition. Chapter 3 introduces the database that we developed for this research. The chapter also explains experimental procedure and methodology in analysing the database including the facial features and classifier that are used for the experiment. Result and analysis from the experiment are discussed in Chapter 4. Chapter 5 concludes the thesis with a thought on limitations, potential improvements and forthcoming research considerations.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This part of the thesis explains the history and theory required for better understand the purpose of this thesis. Firstly, Section 2.2 explains the differences between spontaneous and non-spontaneous facial expressions. It also includes the groundwork and benefits of studying spontaneous facial expressions. Thereafter, existing spontaneous facial expression database and its analysis were stated. Next, it explained the history of researches including the existing approaches for facial expressions analysis. Then, current studies of ethnic classification and explained the differences between ethnicity using facial expressions were discussed. After that, the comparation of existing ethnicity facial database and their analysis was made. Lastly, Section 2.7 explained the theory of Facial Action Coding System (FACS) and its application in the deformation of facial expressions.

2.2 Existing Works on Facial Expression

Since spontaneous facial expression recognition was naturally expected a greater challenge than posed expression (Wan & Aggarwal, 2014), deep studies have been done to obtained better recognition results. Adaptive Online Metric Learning (AOML) algorithm, a new metric space of facial expression was introduced by Wan and Aggarwal (2014). The precision of annotation is defined by assigning a gold standard label to each expression based on majority voting of annotators. The experiment was done using existing Moving Faces and People (MFP) dataset (O'Toole et al., 2005). Landmark points were automatically placed using Constrained Local Model (CLM) (Saragih et al., 2009) before proceed to feature extraction using Gabor filters. The recognition process was then proceed using a Robust Metric Learning approach (RobustML) and Adaptive Online Metric Learning algorithm, which helped obtaining faster convergence by adjusting the gradient adequate step size and decreasing computational tasks via substitution of eigen-decomposition based positive semidefinite (p.s.d) with much simple two-case function. The recognition accuracy method was then compared with Euclidean distance metric (EUC), Isomap (Tenenbaum et al., 2000), Locally Linear Embedding (LLE) (Saul & Roweis, 2000), SVM, Latent-Dynamic Conditional Random Field (LDCRF) (Roweis & Saul, 2000) and non-robust version of metric learning (nrML). From the comparison, RobustML and nrML were the most outstanding in differentiating the "neutral", "fear" and "sad" expression, whereas LDCRF resulted the highest recognition accuracy.

Research by El Meguid et al. (2014) deliberated the framework for fully automated spontaneous facial expression from videos using PittPatt face detector (Pittsburgh, 2011) and Random Forest tree paired with SVM classifier. The experiment done using 7 expressions of Binghamton University 3D Facial Expression (BU-3DFE) database for training and tested with few labelled spontaneous databases (BU-4DFE (Yin et al., 2008), FEED (Wallhoff, 2006) and AFEW database (Dhall et al., 2011). For this experiment, BU-4DFE obtained the best results because it originated in the same set up as the training database. Most of the confusions from the experiment were oriented by the low-intensity of the expressions. FEED database obtained lower result from incorrect frame labelling for certain video clip and also from low intensity expressions. AFEW database that contained acted spontaneous expressions resulting lowest result among all. This may due to the posed expression from actors may not represent actual human emotion.